CLOUGH

BOSTON UNIVERSITY

GRADUATE SCHOOL

Thesis

MATRICES

By
Lucy Isabel Clough
(A.B., Colby College, 1913)

submitted in partial fulfillment of the
requirements for the degree of
Master of Arts

1931

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

Scr

p6861

upstairs
378.744
BO
A.M. 1931
d

Outline

Pages 6 - 8

I. Introduction

A. Historical Summary.

B. Notations in Use.

II. Preliminary Definitions and Theorems. 8 - 12

A. Preliminary Definitions Concerning Determinants.

1. Minor of element of determinant.

2. Complement of k-rowed minor of determinant.

3. Algebraic Complement of k-rowed minor of determinant.

4. Principal minor . " "

5. Co-factor of element . " "

6. Bordered determinant.

7. Adjoint of determinant.

B. Statement of Certain Theorems of Determinants.

1. Theorem concerning interchange of corresponding rows and columns.

2. Theorem concerning interchange of two rows (or of two columns).

3. Theorem concerning elements of two rows identical (or 2 columns).

4. Theorem concerning elements of two rows proportional (or 2 columns).

5. Theorem concerning multiplication of elements of row (or column) by constant.

6. Theorem concerning binomial elements of row (or column).

7. Theorem concerning addition to elements of row (or column) of product of elements of another row (or column) by same constant.

8. La Place Expansion of determinant.

9. Theorem concerning product of two determinants of n^{th} order.

Pages

10. Cramer's Rule.

11. Three theorems concerning value of bordered determinants.

12. Three theorems concerning adjoints of determinants.

13. Two theorems concerning co-factors of determinants.

III. Discussion of Nature of Matrix and Its Properties. 12 - 19

A. Definition of Matrix.

B. Properties of Matrix.

C. Theorems Concerning Rank of Matrix.

1. (Preliminary) A necessary and sufficient condition for the linear dependence of M sets of N constants each is that, all the m-rowed determinants of the matrix a shall vanish, (M ≥ N).

$$\underline{a} = \begin{Vmatrix} X_1^{(1)} & X_2^{(1)} & \cdots & X_N^{(1)} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ X_1^{[M]} & X_2^{[M]} & \cdots & X_N^{[M]} \end{Vmatrix}$$

2. (Preliminary) If a set of points, finite or infinite in number, have the property that k points can be found among them on which any other point in the set is linearly dependent, then any (k+1) points of the set will be linearly dependent.

3. If in a given matrix a certain r-rowed determinant is not zero, and all the (r+1) - rowed determinants, of which this r-rowed determinant is a first minor, are zero, then all the (r+1)-rowed determinants of the matrix are zero.

4. The rank of a matrix is not affected by any of the three elementary transformations.

5. Two equivalent matrices have the same rank.

6. Two matrices with m rows and n columns are equivalent if they have the same rank.

IV. Symmetrical Matrices. 20 - 22

A. Definition of Symmetrical Matrix.

Pages

B. Three Theorems Concerning the Rank of a Symmetrical Matrix

1. If an r-rowed principal minor M_r of the symmetrical matrix $\underline{a}$ is not zero, while all the principal minors obtained by adding one row and the same column, and also those obtained by adding two rows and the same two columns to M_r, are zero, then the rank of $\underline{a}$ is r

2. If all the (r+1)-rowed principal minors of the symmetrical matrix $\underline{a}$ are zero, and also all the (r+2)-rowed principal minors, then the rank of $\underline{a}$ is r, or less.

3. If the rank of the symmetrical matrix $\underline{a}$ is $r > 0$, there is at least one r-rowed principal minor which is not zero.

V. The Algebra of Matrices. 23 - 26

A. Matrices as Complex Quantities.

B. Necessary Definitions.

1. Zero matrix.

2. Equality of matrices.

3. Sum (or difference) of two matrices.

4. Product of matrix and a scalar.

5. Singular matrix.

6. Divisor of zero.

7. Product of two square matrices.

8. Self-evident Laws.

C. Theorems Concerning Multiplication of Matrices

1. Not in general commutative.

2. Associative.

3. Distributive.

D. Theorems concerning Equality.

1. If it is possible to pass from $\underline{a}$ to $\underline{b}$ by an elementary transformation there exists a non-singular matrix $\underline{c}$ such that $\underline{ac} = b$ or a non-singular matrix $\underline{d}$ such that $\underline{da} = b$.

Pages

VI. The further Development of the Algebra of Matrices. 27 - 29

A. Two Theorems Concerning Unit Matrix I .

B. Laws of Indices for Matrices .

C. Theorem Concerning Division of Matrices .

D. Theorem: any k-rowed determinant of matrix $\underline{ab}$ is equal to an aggregate of k-rowed determinants of $\underline{b}$, each multiplied into a polynomial in the a's and likewise for k-rowed determinants of $\underline{a}$ multiplied into a polynomial in b's.

E. Two Theorems Concerning the Rank of the Product of Two Matrices.

1. The rank of the product of two matrices cannot exceed the rank of either factor.

2. If a matrix of rank r is multiplied in either order by a non-singular matrix, the rank of the product is also r.

F. Two Theorems on the Equivalence of Matrices .

VII. λ - Matrices. 30 - 37

A. Definition of λ matrix: Properties of λ - Matrix .

B. Theorems Concerning λ - Matrices:

1. If $\underline{a}$ and $\underline{b}$ are equivalent λ - matrices of rank r and if $D_i(\lambda)$ is the greatest common divisor of the i-rowed determinants of $\underline{a}$, then it is also the greatest common divisor of the i-rowed determinants of $\underline{b}$

2. Every λ-matrix of the nth order and of rank r can be reduced by elementary transformations to the normal form,

$$\begin{Vmatrix} E_1(\lambda) & 0 & \cdots & 0 & 0 \\ 0 & E_2(\lambda) & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & E_r(\lambda) & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & 0 \end{Vmatrix}$$

Pages

3. The greatest common divisor of the i-rowed determinants of a λ-matrix of rank r, $(1 \leqq r)$ is $D_i(\lambda) = E_1(\lambda), E_2(\lambda) \cdots E_i(\lambda)$ where the E's are the elements of the normal form to which the given matrix is equivalent.

4. A necessary and sufficient condition for the equivalence of two λ-matrices of the nth order is that they have the same rank r and that for every value of i from 1 to r inclusive, the i-rowed determinants of one have the same greatest common divisor as the i-rowed determinants of the other.

VIII. Summary. 38

A. Three Types of Matrices Considered.

1. Ordinary matrix of system of M N quantities arranged in rectangular array of M rows and n columns

2. Symmetric matrices in which the pairs of equal elements are symmetric with respect to the leading diagonal.

3. λ-Matrices in which the elements are polynomials in a single variable λ.

B. Three Main Groups of Theorems.

1. Theorems Concerning the rank of a matrix.

2. Theorems Concerning the equivalence of two matrices.

3. Theorems Concerning the laws of algebra for matrices

IX. Bibliography. 39 - 41

Matrices

Historical Summary

Determinants and their logical development, matrices, are a comparatively recent phase of mathematics. We find the former discussed for the first time by Leibniz in a letter to L'Hospital in April 1693. He did not, however, develop the new theory to any great extent. The next contributor of note was Cramer, who, in 1750 in his studies of the solution of sets of linear equations, evolved the famous Cramer's Rule. After this, determinants were recognized as rather valuable tools and Gauss, Jacobi, Laplace and several others extended the number of theorems in this field. From such beginnings, Cayley in 1843-1845 became the creator of the theory of matrices in his Calculus of Matrices. In connection with their work on linear transformations, invariants, quadratic forms, quaternions and vectors we find many later writers extending the theorems of matrices. Among the best known of these authors are Frobenius, Grassman, Krönecker and Weierstrass. Quite recently in 1915-1925 Dr. C. E. Cullis of the University of Calcutta has developed a most extensive discussion of this whole subject in three huge volumes concerned entirely with matrices and determinoids, a further extension of the original form. Of the works of the modern mathematicians who have written extensively on matrices, that of Dr. Maxime Bôcher in his "Introduction to Higher Algebra" seems to me the most easily understood, the most logical and the most compact. I have, therefore, followed his development of the topic very closely. I was, in fact, compelled to do so, since both Drs. Whitehead and Cullis have built up a terminology so extensive that its explanation alone would require a paper equal in length to this article while Dickson has so involved his discussion of

matrices with the theory of linear transformations, and with invariant and quadratic forms, that it is difficult to explain his theorems without depending too largely on these other phases of mathematics.

There are several notations in use. Dr. Cullis uses the form $A = (a)_m^n$ to stand for a matrix $\underline{a}$ = a series of m, n quantities where n represents the rows and m the columns. Others use the simple brace on the left of quantities as

$$\begin{cases} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \end{cases}$$

But the more widely known is that given by Bôcher, two parallel straight lines on either side the quantities contained:

$$\underline{a} = \begin{Vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{Vmatrix}$$

Since the proofs of the theorems concerning matrices are to a very great extent dependent on the properties of the determinants of such matrices, the following definitions and theorems are listed here as references. Their proofs may be found in any text on determinants.

II. Definitions:

1) If the row and column containing a given element be removed, the determinant of one less order, remaining, is called the minor of that element.

2) If D is a determinant of the nth order, and M one of its k-rowed minors, then the (n-k)-rowed minor N obtained by striking out from D all the rows and columns represented in M is called the complement of M.

3) If M is the m-rowed minor of the determinant D in which the rows $k_1 - - - k_m$, and the columns $l_1 - - - l_m$, are represented, then the algebraic complement of M is defined by the equation,

algebraic complement of M = $(-1)^{k_1 - - - k_M + l_1 - - - l_M}$ [Complement of M]

4) By a principal minor of a determinant D is understood a minor obtained by striking out from D the same rows as colums.

5) If a_{ij} be the element in the ith row and jth column of a determinant, and D_{ij} represents its corresponding first minor, $A_{ij} = (-1)^{i + - - + j} D_{ij}$.

6) If to a determinant of the nth order be added one, or more, rows and the same number of columns of n quantities each, and if the vacant corner is filled in with zeros, the determinant thus obtained is called a bordered determinant.

7) If in any determinant D, each element be replaced by its cofactor, the resulting determinant D' is called the adjoint determinant.

B. Preliminary Theorems on Determinants:

1) If in any determinant the corresponding rows and columns be interchanged, the value of the determinant remains the same.

2) If the elements of a row (or column) be multiplied by a constant, the effect is to multiply the determinant by that constant.

3) If the elements of two rows (or columns) of a determinant are identically the same, the determinant vanishes.

4) If the elements of two rows (or columns) of a determinant are proportional, the determinant vanishes.

5) If the corresponding elements of two rows (or of two columns) be interchanged, the sign of the determinant is reversed, or remains the same according as $(-1)^{i+j}$ --+ M is negative or positive, if i and M are the rows (or columns) to be interchanged.

6) If the elements of a given row (or column) be multiplied by a constant and added to the corresponding elements of another row (or column) the value of the determinant is unchanged.

7) If each element of a given row (or column) of a determinant is a binomial, the determinant may be expressed as the sum of two determinants.

8) If from any determinant D, any K rows (or columns) be selected and from these rows (or columns) all possible kth order determinants be obtained; and if each of these is multiplied by its complementary minor, and to each product is prefixed a + or- according as $(i_1 + i_2 + \cdots + i_k) + (j_1 + j_2 + \cdots + j_k)$ is even, or odd, the algebraic sum of all such products is the value of the given determinant. This is called the Laplace Expansion. j and i represent the columns and rows concerned.

9) The product of two determinants of the nth order may be expressed as a determinant of the nth order, in which the element which lies in the ith row and jth column is obtained by multiplying each element of the ith row, of the first factor, by the corresponding element of the jth column of the second factor and adding the results.

10) Cramer's Rule: If in the equations,

$$a_{11}X_1 + \cdots + a_{1N}X_N = K$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$a_{N1}X_1 + \cdots + a_{NN}X_N = K_N$$

the determinant

$$a = \begin{vmatrix} a_{11} & \cdots & a_{1N} \\ \cdots & \cdots & \cdots \\ a_{N1} & \cdots & a_{NN} \end{vmatrix} \neq 0,$$

the equations have one and only one solution which is:

$$X_1 = \frac{a_1}{a} \; ; \; X_2 = \frac{a_2}{a} \; ; \; X_N = \frac{a_N}{a}$$ if

a_i is the n-rowed determinant obtained from "a" by replacing the elements of the ith column by elements $k_1, k_2, \cdots k_N$.

11) If a determinant of the nth order is bordered with n rows, and n columns, the resulting determinant has a value which depends only on the bordering quantities.

12) If a determinant of the nth order is bordered with more than n rows and n columns, the resulting determinant always has the value zero.

13) If D' is the adjoint of determinant D and M and M' are corresponding M-rowed minors of D and D' respectively, then M' · ·

equals the product of D^{N-1} by the algebraic complement of α.

14) If a_{iJ} is an element of a determinant D of the nth order, and if α_{iJ} is the co-factor of the corresponding element A_{iJ} of the adjoint of D, then

$$A_{iJ} = D^{N-2} \alpha_{iJ}$$

15) If D is any determinant of the nth order and D' is its adjoint,

$$D' = D^{N-1}$$

16) If the elements of a row (or column) of a determinant be multiplied by the co-factors of another row (or column), and added, the result vanishes.

17) If D is any determinant and S is the second minor obtained from it by striking out the ith and kth rows, and its jth and lth columns, and if by A_{iJ} is denoted the co-factor of the element which stands in the ith row and Jth column of D, then

$$\begin{vmatrix} A_{ij} & A_{il} \\ A_{kj} & A_{kl} \end{vmatrix} = (-1)^{i+j+k+l} D S.$$

Part III.

A. Nature of Matrix

B. Properties of Matrix

All these prelimiary definitions and theorems of determinants being granted, it is now possible to begin the discussion of matrices, for although matrices are not determinants, they are a development of the latter and borrow very freely from that subject for their needs.

A matrix, then, is defined as a system of M N quantities arranged in a rectangular array of M rows and N columns. M may, or may not equal N. Should M equal N we have a square matrix of order N. Even if, however, M does equal N, the matrix thus formed is not a determinant for a determinant is either a scalar or a polynomial while a matrix is not a quantity at all but a system of quantities. Bôcher, illustrates this differences in the following way. If the rows and columns of a determinant be interchanged, the value of the determinant is in no way affected. (See (1) section II). In the case of a matrix, on the contrary, such an interchange gives a wholly new matrix which is called the conjugate of the first square matrix.

Various properties of a matrix are to be established. A determinant is said to determine a square matrix, the matrix of the determinant. Conversely a square matrix determines a determinant which we term the determinant of the matrix. Every matrix has several determinants, which are to be obtained from the square matrices formed by striking out certain rows (or columns) or both. All these are named determinants of the original matrix, and they may be of any order from one to the smaller of the two numbers, M and N. Some of them, naturally, must be of the form of the bordered determinants of (12) Section II and so must equal zero.

A matrix is said to be of rank r if it contains at least one r-rowed determinant which is not equal to zero, while all the (r+1)-rowed determinants or all determinants of order higher than r, are equal to zero.

$$\underline{a} = \begin{Vmatrix} 14 & 12 & 6 & 3 & 2 \\ 6 & 104 & 21 & 9 & 17 \\ 7 & 6 & 3 & 4 & 1 \\ 35 & 30 & 15 & 20 & 5 \end{Vmatrix}$$

Here matrix a must be of second order, for it can be shown that all fourth order and all third order determinants must each equal zero, since they must all contain two proportional rows. One second order determinant not equal to zero can be found. Obviously, then, a matrix must be of rank zero if all its elements are zero.

To find the rank of a matrix, then, we must find one r-rowed determinant not equal to zero, and next show that all the (r+1)-rowed determinants are equal to zero. In most cases this would be, necessarily, a lengthy proceeding and therefore several methods of shortening this procedure have been devised. But before we can prove the first theorem in this direction, two preliminary theorems must be established, and for these we must turn to the theory of linear dependence.

C. M sets of N constants each $(x_1^{[i]}, x_2^{[i]} - - - x_M^{[i]})$ if $(i = 1, 2 - - - M)$ are said to be linearly dependent if M constants not all zero exist such that $C_1 x'_j + C_2 x''_j + - - + C_M x^{[M]}_j = 0$, $(j = 1, 2, - - - N)$.

If $M \leqq N$, the following theorem may be proved.

Preliminary 1: A necessary and sufficient condition for the linear dependence of M sets of N constants each, (when $M \geqq N$) is that all the M-rowed determinants of the matrix $\begin{Vmatrix} x'_1 & x'_2 - - - - - x'_N \\ \vdots & \vdots \\ x^{[M]}_1 & x^{[M]}_2 - - - - x^{[M]}_N \end{Vmatrix}$ should vanish.

Necessary Condition: Since the M sets of constants are given linearly dependent, by the preceding definition one row may be expressed as a linear combination of the others. Then we may reduce any of the M-rowed determinants to zero by subtracting from the elements of this row the corresponding elements of the other rows multiplied by suitable constants. The elements of this row can thus be reduced to zero and so the M-rowed determinant vanished by theorem 4 Section II.

Sufficient condition: If all M-rowed determinants of the matrix vanish, then the M sets of N constants each are linearly dependent. If all the M-rowed determinants vanish, we still may assume that the rank of the matrix is $r > 0$, for otherwise each element of the matrix must equal zero. Since the rank of the matrix is r, there must be at least one r-rowed non-vanishing determinant. By a sufficient number of changes of rows, this non-vanishing r-rowed determinant may be shifted to the upper left hand corner. We have, then, a non-vanishing r-rowed determinant in the upper left hand corner. Let C_1, C_2, C_{r+1} be the co-factors of the (r+1)-rowed determinant in upper left hand corner of the matrix corresponding to the elements of the last column. Now from our definition of the rank of matrix = r, the (r+1)-rowed determinants must vanish. But the sum of the products of elements of any column of a determinant by the co-factors of the corresponding elements of another column is zero. (See Section II, Theorem 16.) Therefore from our definition of linear dependence, the first (r+1) sets of constants are linearly dependent since c_{r+1} is not zero, being the r-rowed determinant in the upper left hand corner. If $M > N$, by adding to each set of n constants (m-n) zeros, we may reduce the theorem to one just proven. So M sets of N constants each are always linearly dependent if $M > N$.

Preliminary 2: If a set of points, finite or infinite in number, has the property that k points can be found among them on which every other point of the set is linearly dependent, then any (k+1) points of the set will be linearly dependent.

Before we go on to the steps of the proof of this second theorem, we must again go back to certain preliminary definitions of linear independence. The N sets of N constants each with which we had to deal in the preliminary definition of linear dependence and in the preceding proof may be considered as the n homogeneous coördinates of m points in a space of (n-1) dimensions, provided that not all the constants are zero. If these constants are so considered, we may then speak of the linear dependence of the points, whose coördinates, the series of N constants have then become.

In preliminary theorem 2, then, $(p_1 - - - p_n)$ represent the homogeneous coördinates of a point in a space of (N-1) dimensions. Let $(p'_1 - - - p'_n)$, $(p''_1 - - - p''_n)$, - - - - $(p^{[k]}_1 - - - p^{[k]}_n)$ represent the k points on which every other point of the set is linearly dependent and let $(P'_1 - - - P'_n)$, $(P''_1 - - - P''_n)$, - - - $(P^{[k+1]}_1 - - - P^{[k+1]}_n)$ represent any (k+1) points of the set.

If $P^{[i]}_1 - - - P^{[i]}_N$ is any one of these points, it is by hypothesis linearly dependent on any one of the k points so there must exist a set of k constants not all zero $c_1^{[i]}, c_2^{[i]} - - - c_K^{[i]}$, such that $(P_1^{[i]} - - - P_N^{[i]})$ may be written:

$$(i = 1, 2 - - - N) \qquad (1) \qquad \begin{cases} P_1^{[i]} = c_1^{[i]} p'_1 + c_2^{[i]} p''_1 + c_3^{[i]} p'''_1 + - - - + c_K^{[i]} p_1^{[K]} \\ - \\ - \\ P_N^{[i]} = c_1^{[i]} p'_N + c_2^{[i]} p''_N + c_3^{[i]} p'''_N + - - - + c_K^{[i]} p_N^{[K]}. \end{cases}$$

This will also be true if $(P_1^{[i]} - - - P_N^{[i]})$ is one of the K points. If we

are to prove, now, that any (K+1) points are to be linearly dependent, it must be shown that there are (K+1) constants, $C_1, C_2 - - - C_{K+1}$, not all zero, such that $C_1P'_j + C_2P''_j + C_3P_j + - - + C_{K+1}P^{[k+1]}_j = 0$.

$$(j = 1, 2, - - - N)$$

If we substitute for P'_j, etc. the values found in (1) we have

$$(C_1c'_1 + C_2c''_1 - - - - C_{K+1}c^{[k+1]}_1)p'_1 + - - - - - - -$$

$$- +$$

$$(C_1c'_k + C_2c''_k - - - - C_{K+1}c^{[k+1]}_k)p = 0$$

Now these equations will hold true if

$$C_1c'_1 + C_2c''_1 + C_3c'''_1 + - - + C_{K+1}c^{[k+1]}_1 = 0$$

$$- = 0$$

$$- + = 0$$

$$C_1c'_k + C_2c'_k + - - - - - - - + C_{K+1}c^{[k+1]}_k = 0 .$$

But this system of equations is one in which there are fewer equations than there are unknowns, so there must always exist a set of C's not all zero which satisfy the equation. Therefore, by our first definition of linear dependence, the (k+1) points are linearly dependent.

It is now possible to prove the following proposition which materially lessens the steps necessary to determine the rank of a matrix.

Theorem 3: If in a given matrix a certain r-rowed determinant is not zero, and all the (r+1)-rowed determinants of which this r-rowed determinant is a first minor are zero, then all the (r+1)-rowed determinants of the matrix are zero.

Proof: This non-vanishing r-rowed determinant may, we know, be shifted to the upper left hand corner of the matrix. By the first of

the theorems just proved, since the (r+1)-rowed determinants of this r-rowed determinant are zero by hypothesis, the (r+1) sets of constants are linearly dependent. Now the r sets of constants which are in the first r rows of the matrix are not linearly dependent, but linearly independent, since the r-rowed determinant does not vanish. Hence the (r+1)th row must be linearly dependent on the first r-rows. Similarly it may be shown that each of the succeeding rows are linearly dependent on the first r rows. By the second theorem just proved, any (r+1) rows are then linearly dependent. Consequently all (r+1)-rowed determinants of the matrix must be zero (by theorem 1) and therefore theorem 3 is proved.

More practical even than this method is that of simplifying the form of a matrix without altering its rank. The so-called elementary transformations are:

1) The interchange of two rows (or two columns);

2) The multiplication of each element of a row (or column) by the same constant not zero;

3) The addition to the elements of one row (or column) of the products of the corresponding elements of another row (or column) by one and the same constant.

If we can prove that such simplification does not change the rank of a matrix, the process of determining the rank of a matrix becomes much shorter. Now the first two of these transformations will have no effect on the rank of any matrix since they do not have any effect on the vanishing of any determinant by B (2) and (5) of Section II. It is necessary, then, only to show that the third transformation does not alter the rank of a matrix.

Proof: Suppose matrix b is formed from matrix a by adding to the elements of some row, the ith, k times the elements of the jth row.

If r is the rank of the matrix $\underline{a}$, all its (r-1)-rowed determinants must be zero from the definition of the rank of a matrix. Now some of these (r-1)-rowed determinants will not be changed by this addition, those containing both the ith and jth rows, and those which do not contain the ith row. The rest of these (r-1)-rowed determinants can be split up into the form $A \pm kB$ where A and B are (r-1)-rowed determinants of $\underline{a}$ by theorem 7, Section II. Since by hypothesis all the (r-1)-rowed determinants of $\underline{a}$ are 0, and since A and B are (r-1)-rowed determinants, $A \pm kB$ must equal zero. Hence all the (r-1)-rowed determinants resulting from the transformation are zero. Therefore the rank of $\underline{a}$ is not increased. Nor can the rank of $\underline{a}$ have been decreased, for if the matrix $\underline{b}$ secured by the transformation be of lower rank than $\underline{a}$, the transformation back to $\underline{a}$ would be a process similar to the one just described. It has just been shown that such a process does not increase the rank, so $\underline{b}$ cannot be of lower rank than $\underline{a}$. We have therefore proved

Theorem 4: The rank of any matrix is not changed by any one of the three elementary transformations.

By definition, two matrices are equivalent if it is possible to pass from one to the other by any finite number of elementary transformations. We may state, then, as a direct result of theorem 4,

Theorem 5: Two equivalent matrices have the same rank.

Any matrix of rank r can be reduced by means of elementary transformations to a form where the element in the ith row and ith column is 1, when $i \leq r$, while all the other elements of the matrix are zero. This is true for a matrix of rank r, since the only determinants not

zero must be r-rowed, that is, the matrix may be reduced to one of r rows. If $i \leqq r$, by elementary transformation 3, the other rows (except the ith) may be reduced to zero, by a finite number of transformations. Then by a finite number of transformations the column may also be reduced to zero with the exception of the element in the ith column. The following proposition is the obvious result of such transformations:

Theorem 6: Two matrices with M rows and N columns are equivalent if they have the same rank. For if they have the same rank and are of same order, each may be reduced to an equivalent matrix where the element in ith row, ith column is 1 if $i \leqq r$.

IV. Symmetric Matrices

A. Definition of Symmetric Matrix

A symmetric matrix, like a symmetric determinant, is a matrix in which the pairs of terms situated symmetrically with respect to the principal diagonal, are equal. To determine the rank of such a matrix, the following three theorems have been discovered.

B. Three Theorems Concerning Rank of Symmetric Matrix.

Theorem 1: If an r-rowed principal minor M_r of the symmetrical matrix $\underline{a}$ is not zero, while all the principal minors obtained by adding one row and the same column, and also those obtained by adding two rows and the same two columns to M_r, are zero, then the rank of $\underline{a}$ is r.

Obviously if we can show that all the principal minors except the r-rowed minors are equal to zero, this theorem resolves itself into that of (3) of section III. As usual, we assume that the r-rowed principal minor not equal to zero has its position in the upper left hand corner, as it may be placed there by a finite number of interchanges, first of rows, then of columns. Let B_{ab} represent the determinant formed by adding the ath row and bth column of $\underline{a}$ to M_r. By hypothesis B_{aa} and B_{bb} would be equal to zero, and $M_r \neq 0$. Let C represent the determinant formed by adding ath and bth columns and ath and bth rows to M_r. Then $C = 0$ by hypothesis, also. If M'_2 be the two-rowed principal minor of the adjoint of C which corresponds to the complement of M_r in C, then by theorem 17 of Section II

$$\begin{vmatrix} B_{aa} & B_{ab} \\ B_{ba} & B_{bb} \end{vmatrix} = M'_2 = C\, M_r = 0$$

Therefore $(B_{aa})\,(B_{bb}) - (B_{ab})^2 = 0$

Since $B_{aa} = 0$ and $B_{bb} = 0$; $B_{ab} = 0$ Thus all the determinants except the

r-rowed principal minors are zero and hence the rank of the symmetrical matrix is r.

Therorem 2: If all the (r+1)-rowed principal minors of a symmetrical matrix $\underline{a}$ are zero, and also all the (r+2)-rowed principal minors, the rank of $\underline{a}$ is r or less.

Proof: Suppose r = 0. Then all the elements of the principal diagonal must equal zero. Hence the 'two'-rowed principal minors must then equal zero. Therefore by theorem 17, Section 11

$$(a_{ii})\,(a_{jj}) - (a_{ij})^2 = 0$$

As $a_{ii} = 0$, and $(a_{jj}) = 0$, a_{ij} must also equal zero. Each element is therefore zero, so the rank of the matrix in this case must be zero, and thus equal to r. Thus the theorem holds true for this special case.

Next suppose the theorem holds good for r= k. If the (k+1)-rowed and (k+2)-rowed principal minors equal zero, then r should equal k or less. Similarly, if the (k+2)- rowed and (k+3)-rowed principal minors are zero, then the rank of this matrix should be less than k-1 by the definition of rank of matrix. If the(k+1)-rowed minors are not zero while (k+2)-rowed and (k+3)-rowed determinants are zero, then the rank of the matrix is exactly k+1 by theorem 1. Therefore the theorem holds true for k+1 if it holds true for r = k. But in the first part of this proof we proved the theorem true for r = 0, so it must hold true for r = 1, 2, and all successive values.

Theorem 3: If the rank of the symmetrical matrix $\underline{a}$ is $r > 0$, there is at least one r-rowed principal minor of $\underline{a}$ which is not equal to zero.

Proof: The (r-1)-rowed principal minors are zero. (By definition of rank.)

If all the r-rowed principal minors also were equal to zero, the rank

of the matrix would have to be less than r by the preceding theorem. But this is contrary to the hypothesis, so there must be at least one r-rowed principal minor which is not equal to zero.

V. The Algebra of Matrices.

A. Matrices as Complex Quantities with the Necessary Definitions.

In the definition of a matrix as a rectangular array of M N quantities, no distinction was made as to whether these systems of quantities were composed of real or complex quantities. The definitions and theorems of section V, however, are discussed from the viewpoint of matrices as complex and made up of certain components, real or unreal. Dr. Bôcher states the following definitions as necessary postulates:

1) A matrix is said to be zero, when and only when all of its elements are zero.

2) Two matrices are said to be equal when and only when they have the same number of rows and of columns, and every element of one is equal to the corresponding element of the other.

3) By the sum (or difference) of two matrices, of m rows and n columns each, is meant a matrix of m rows and n columns, each of whose elements is the sum (or difference) of the corresponding elements of the given matrices.

4) The product of a matrix and a scalar is a matrix each of whose elements is k times the corresponding element of the original matrix, if k equals the given scalar.

5) A square matrix is said to be singular if its determinant is zero.

6) A matrix, $\underline{a}$, is said to be a divisor of zero, if a matrix, $\underline{b} \neq 0$, exists such that $\underline{ab} = 0$, or $\underline{ba} = 0$.

7) The product $\underline{ab}$ of 2 square matrices of the nth order is a square matrix of the nth order, in which the element which lies in the ith row and jth column is obtained by multiplying each element of the ith row of $\underline{a}$ by the corresponding element of the jth row of $\underline{b}$ and adding the result.

This latter definition is given for square matrices only, but may be applied to matrices of M rows and N columns where $M \neq N$, if we consider such a matrix equivalent to a square matrix of order equal to the larger of the two integers M and N, in which the lacking rows or columns are filled in with zeros. We may therefore consider all matrices as square, then, in our future discussions, if so completed.

The following statement follows naturally then from the foregoing definitions:

8) All the laws of ordinary algebra hold for the addition, or subtraction of matrices, and their multiplication by scalars.

If then, $\underline{a}$, $\underline{b}$, and $\underline{c}$ are matrices and M and N are scalars,

$$\underline{a} + \underline{b} = \underline{b} + \underline{a},$$

$$\underline{a} + (\underline{b} + \underline{c}) = (\underline{a} + \underline{b}) + \underline{c},$$

$$M\underline{a} + M\underline{b} = M(\underline{a} + \underline{b}),$$

$$M\underline{a} + N\underline{a} = (M + N)\underline{a}$$

B. Theorems Concerning Multiplication of Matrices.

The first theorem of this section on the algebra of matrices is:

Theorem 1: The multiplication of matrices is not in general commutative. Let a_{ij} and b_{ij} represent elements in the ith row and jth column of two matrices $\underline{a}$ and $\underline{b}$ to be multiplied. By definition 7, above, the element in ith row, jth column of their product $\underline{ab}$ will be

$$a_{i1}b_{1j} + a_{i2}b_{2j} + - - - - - - + a_{iN}b_{Nj}$$

If the multiplication were in the other order, the corresponding element of $\underline{ba}$ would be

$$a_{1j}b_{i1} + a_{2j}b_{i2} + - - - - - - + a_{Nj}b_{iN}$$

These do not always have to be equal, so in general $ab \neq ba$.

Theorem 2: The multiplication of matrices is associative:

Given in this a third matrix $\underline{c}$ whose element $(ij) = C_{ij}$ and also given two matrices $\underline{a}$ and $\underline{b}$ as above with elements as given in (1). Then element (ij) of $(\underline{ab})\underline{c}$ is $[1]$; $(a_{i1}b_{11} + a_{i2}b_{21} + - - - + a_{iN}b_{N1})c_{1j} + (a_{i1}b_{12} + a_{i2}b_{22} + - - - + a_{iN}b_{N2})c_{2j} + - - - + (a_{i1}b_{1N} + a_{i2}b_{2N} + - - - + a_{iN}b_{NN})c_{Nj}$. But the element (ij) of $\underline{a}(\underline{bc})$ is $[2]$; $a_{i1}(b_{11}c_{1j} + b_{12}c_{2j} + - - + b_{1N}c_{Nj}) + a_{i2}(b_{21}c_{1j} + b_{22}c_{2j} + - - - + b_{2N}c_{Nj}) + - - - + a_{iN}(b_{N1}c_{1j} + b_{N2}c_{2j} + - - + b_{NN}c_{Nj})$. $[1] = [2]$. Hence $(\underline{ab})\underline{c} = \underline{a}(\underline{bc})$.

Theorem 3: The multiplication of matrices is distributive, that is,

$$\underline{a}(\underline{b} + \underline{c}) = \underline{ab} + \underline{ac}$$

The element (i,j) of matrix $\underline{a}(\underline{b} + \underline{c})$ equals the sum of the elements (i,j) of matrices $\underline{ab} + \underline{ac}$, obviously.

Theorem 4: If it is possible to pass from $\underline{a}$ to $\underline{b}$ by an elementary transformation, there exists a non - singular matrix $\underline{c}$ such that

$$\underline{ac} = \underline{b},$$

or a non-singular matrix $\underline{d}$ such that

$$\underline{da} = \underline{b}$$

Proof: Given matrix $\underline{a}$ and matrix $\underline{b}$ derived by an elementary transformation from $\underline{a}$.

Case 1: Where the elementary transformation consists of the interchange of two rows (or two columns). Here $\underline{b}$ is of the form secured by a multiplication of this type:

If $\underline{a} = \begin{Vmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ a_{31} & a_{32} & \cdots & a_{3N} \\ \cdots & \cdots & \cdots & \cdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{Vmatrix}$ and, $\underline{c} = \begin{Vmatrix} 0 & 1 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots \\ 0 & 0 & 0 & 1 \cdots \\ 0 & 0 & 0 & 0 \cdots 1 \end{Vmatrix}$, then $\underline{ac} = b$,

for such a multiplication would give us a matrix $\underline{b}$ in which the first two columns of $\underline{a}$ have been interchanged. $\underline{c}$ is not zero, so this case holds good: $\underline{ac} = \underline{b}$. $\underline{da} = \underline{b}$ may be similarly derived by the multiplication of the two matrices: Here $\underline{b}$ is secured from $\underline{a}$ by the interchange of two rows.

$$\underline{d} = \begin{Vmatrix} 0 & 1 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots \\ 0 & 0 & 0 & 1 \cdots \\ 0 & 0 & 0 & 0 \cdots 1 \end{Vmatrix} \quad \text{and} \quad \underline{a} = \begin{Vmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ a_{31} & a_{32} & \cdots & a_{3N} \\ \cdots & \cdots & \cdots & \cdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{Vmatrix}$$

Case 2: Where the transformation is the multiplication of a row (or column) by a constant a similar method is possible where k instead of 1 is the element of the leading diagonal of $\underline{c}$ or $\underline{d}$.

Case 3: When the transformation is of the form of the addition to the elements of one row (or column) of the product of elements of another row (or column) by a constant, the leading diagonal of $\underline{c}$ or $\underline{d}$ may be a binomial.

VI. Further Development of the Algebra of Matrices.

A. Two Preliminary Theorems

Before going on to the theorems on the divisions of matrices and the law of indices, we must turn back for some definitions on linear transformations, or change of variables. The first of these definitions is:

If $x_1 - - - x_n$ are the original variables, and $x'_1 - - - x'_n$ are the new variables, the formulae of these transformations are

$$x'_1 = a_{11}x_1 + - - - + a_{1N}x_N,$$
$$\sim \quad \sim \quad - \quad - \quad - \quad \sim$$
$$x'_N = a_{N1}x_1 + - - - + a_{NN}x_N .$$

The square matrix $\underline{a} = \left\| \begin{matrix} a_{11} - - - a_{1N} \\ - - - - - - - \\ - - - - - - - \\ a_{N1} - - - a_{NN} \end{matrix} \right\|$ is called the matrix of the transformation. The determinant of this matrix is the determinant of the transformation. The reverse of this transformation:

$$x_1 = \frac{A_{11}}{a}x' + - - + \frac{A_{1N}x'_N}{a}$$
$$\sim \quad \sim \quad - \quad - \quad - \quad \sim$$
$$x_N = \frac{A_{N1}}{a}x' + - - + \frac{A_{NN}}{a}x'_N \quad ,$$

where A_{11} etc. are co-factors of $a_{11} - - - - a_{NN}$, is called the inverse of the transformation. a in this case is the determinant of the matrix. It can, of course, exist only if $a \neq 0$, otherwise it is indeterminate. So if a does equal zero, the linear transformation is given the name of a singular transformation.

Still another definition is needed before we can go on to our next theorem. If $x'_1 = x_1$; $x'_2 = x_2$; $x'_N = x_N$ be a special linear transformation whose matrix is $\left\| \begin{matrix} 1 & 0 & - - - & 0 \\ 0 & 1 & 0 & - - - 0 \\ 0 & 0 & 1 & 0 - - 0 \\ 0 & 0 & 0 & 1 - - - 0 \\ - & - & - & - - - - - \\ - & - & - & - - - - - \\ 0 & 0 & 0 & 0 - - 1 \end{matrix} \right\|$ This is called the identical transformation.

Theorem 1: If $\underline{a}$ is any matrix $I\underline{a} = \underline{a}I = \underline{a}$.

A linear transformation of which $\underline{a}$ is matrix will not be changed by being either followed, or preceded by the identical transformation I.

Theorem 2: If $\underline{a} = \begin{Vmatrix} a_{11} & - & - & - & a_{1N} \\ - & - & - & - & - \\ - & - & - & - & - \\ a_{N1} & - & - & - & a_{NN} \end{Vmatrix}$ is a non-singular matrix of determinant a and if A_{ij} denotes co-factors of elements of $\underline{a}$, the matrix $\begin{Vmatrix} \frac{A_{11}}{a} & - & - & - & \frac{A_{N1}}{a} \\ - & - & - & - & - \\ - & - & - & - & - \\ \frac{A_{1N}}{a} & - & - & - & \frac{A_{NN}}{a} \end{Vmatrix}$ called the inverse of $\underline{a}$ and denoted by $\underline{a}^{-1}$, is a non-singular matrix such that $\underline{a}\underline{a}^{-1} = \underline{a}^{-1}\underline{a} = I$.

This is really a restatement of our definitions of linear transformation, its matrix, its inverse and the special transformation I.

3) The laws of indices hold good in part for matrices. If p is any positive integer and $\underline{a}$ any matrix, $\underline{a}^p$ represents $(\underline{a} \times \underline{a} \times \underline{a})$ taken p times as a factor). If $\underline{a}$ is a non-singular matrix $\underline{a}^{-p} = (\underline{a}^{-1})^p$

$$\left.\begin{aligned} \underline{a}^0 &= 1 \\ \underline{a}^p\underline{a}^q &= \underline{a}^{p+q} \\ (\underline{a}^p)^q &= \underline{a}^{pq} \end{aligned}\right\}$$ also as in algebra.

These rules hold true if p and q are positive integers for all matrices. For non-singular matrices p and q may be any integers.

Now we are ready for the division of matrices. Again as in algebra we conceive of this operation as the reverse of multiplication. We showed in theorem 1, section V, that $\underline{ab} \neq \underline{ba}$ in general. This implies that there must then be two quotients in the division of these two matrices $\underline{a}$ and $\underline{b}$. To avoid this difficulty it is usual to express the result of such division in this form:

$\underline{a} = \underline{bx}$ and $\underline{a} = \underline{yb}$, $\underline{x}$ and $\underline{y}$ being two

different matrices. We state it as follows:

Theorem 4: If $\underline{a}$ is any matrix and $\underline{b}$ is any non-singular matrix, there exists one, and only one matrix, $\underline{x}$, which satisfies the equation $\underline{a} = \underline{bx}$, and one and only one matrix, $\underline{y}$, which satisfies the equation $\underline{a} = \underline{yb}$. These matrices are given by the formulae

$$\underline{x} = \underline{b}^{-1}\underline{a} \ , \ \underline{y} = \underline{a}\underline{b}^{-1} \ .$$

Theorem 5: Any k-rowed determinant of the matrix $\underline{ab}$ is equal to an aggregate of k-rowed determinants of $\underline{b}$ multiplied into a polynomial in the a's, and also to an aggregate of k-rowed determinants of $\underline{a}$ multiplied into a polynomial in the b's

Let $\underline{a} = \begin{Vmatrix} a_{11} & a_{12} & - & - & - & a_{1N} \\ a_{21} & a_{22} & - & - & - & a_{2N} \\ a_{31} & a_{32} & - & - & - & a_{3N} \\ - & - & - & - & - & - \\ a_{N1} & a_{N2} & - & - & - & a_{NN} \end{Vmatrix}$ Let $\underline{b} = \begin{Vmatrix} b_{11} & b_{12} & - & - & - & b_{1N} \\ b_{21} & b_{22} & - & - & - & b_{2N} \\ b_{31} & b_{32} & - & - & - & b_{3N} \\ - & - & - & - & - & - \\ b_{N1} & b_{N2} & - & - & - & b_{NN} \end{Vmatrix}$

Take any k-rowed determinant of $\underline{ab}$

$$\begin{Vmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{1N}b_{N1} & - & - & - & a_{11}b_{1N} + a_{12}b_{2N} + a_{1N}b_{NN} \\ - & - & - & - & - \\ a_{N1}b_{11} + a_{N2}b_{21} + a_{NN}b_{N1} & - & - & - & a_{N1}b_{1N} + a_{N2}b_{2N} + a_{NN}b_{NN} \end{Vmatrix}$$

This may be broken up into a sum of determinants of kth order in such a manner that each column of each determinant will have some one of the b's as a factor. If each of these common factors be taken out there will be left a determinant in the a's. If this does not vanish, it is a k-rowed determinant of $\underline{a}$. The process is identically the same for the b's if some one of the a's is taken as a common factor of each column of each determinant. A determinant in the b's will then be left.

Theorem 6: The rank of the product of two matrices cannot exceed the

rank of either factor. From the proof preceding it is plain that if all the k-rowed determinants of $\underline{a}$ or $\underline{b}$ are zero, the k-rowed determinants of $\underline{ab}$ are also zero. Since the rank of a matrix is given by the highest ordered determinant not equal to zero, the rank of the product cannot exceed the rank of either factor.

Theorem 7: If a matrix of rank r is multiplied in either order by a non-singular matrix the rank of the product is r.

Let $\underline{a}$ and $\underline{b}$ represent two matrices of product $\underline{ab}$. Suppose $\underline{a}$ is non-singular. Let r equal rank of $\underline{b}$ and R = rank of $\underline{ab}$.

$R \leqq r$ (Rank of product cannot exceed rank of either factor).

If $\underline{b} = (\underline{a}^{-1})\underline{ab}$,

$r \leqq R$(Rank of product cannot exceed rank of either factor).

Only remaining possibility is R = r.

Similarly if $\underline{b}$ is non-singular and r is taken as rank of $\underline{a}$ and R = rank of $\underline{ab}$, we can prove R = r.

VII. λ - Matrices

A. Definitions of λ -matrices: Properties of λ -Matrices

With the introduction of this topic we begin to see more clearly the value of matrices as a compact method of expressing mathematical groups, and as an aid in the expression of, and solution of, equations for a λ -matrix is one in which the elements are polynomials in one variable λ, and its determinant is also a polynomial in λ. The definitions given for an ordinary matrix hold true also for a λ-matrix- that is - that for a singular matrix, for the three so-called elementary transformations, the rank of a matrix, and the first rule of equivalence. Again, also, if two λ-matrices are equivalent, they are of the same rank.

B. Theorems concerning λ -matrices.

But there is a second method of proving the equivalence of two λ -matrices, to establish which we must first show that the following statement is true:

If in a λ -matrix a. a polynomial, $\phi(\lambda)$, is a factor of all the i-rowed determinants of that matrix, it will be a factor of all the i-rowed determinants of every λ -matrix which can be derived from a by means of elementary transformations.

This we can see is obviously the case for the first two types of transformations, for the first of these does not change the determinant at all, and the second multiplies it only be a constant, or scalar. For the third transformation it may be shown by a series of steps identical to those of theorem 4 section III, if $\psi(\lambda)$ is substituted for constant k (see page 18), that the new determinants thus formed will be of the form $A \mp \psi(\lambda) B$ where A and B are i-rowed determinants of a. So we are able to say that $\phi(\lambda)$ is a factor of all i-rowed determinants formed from a by any one of the three elementary transformations. We may now go on to

our first theorem.

Theorem 1: If a and b are equivalent λ-matrices of rank r, and if $D_i(\lambda)$ is the greatest common divisor of the i-rowed determinants of a when $i \leqq r$, then it is also the greatest common divisor of the i-rowed determinants of b.

Proof: We have just proved that if $D_i(\lambda)$ is a factor of the i-rowed determinants of a, it is also a factor of the i-rowed determinants of b. If it were not the highest common factor of b, there would have to exist another higher factor which would also be a factor of i-rowed determinants of a - for the same reason as above. But this is impossible since $D_i(\lambda)$ was given the highest factor of i-rowed determinants of a. Hence $D_i(\lambda)$ must be the highest common factor of i-rowed determinants of b.

Preliminary 1: If the leading element $f(\lambda)$ of a λ-matrix is not zero, and is not a factor of all the other elements, then an equivalent matrix can be formed whose leading element is not identically zero; and this element will be of lower degree than f.

Case 1: When the element $f_1(\lambda)$, not divisible by $f(\lambda)$, is in the first row. Let j be the column in which $f_1(\lambda)$ is found. Since $f_1(\lambda)$ is not divisible by $f(\lambda)$, there must be a remainder which we will represent by r, and a quotient q. So $f_1(\lambda) = q(\lambda) f(\lambda) + r(\lambda)$. Now to each of the elements of the jth column may be added $-q(\lambda)$ times the first column. This will give us an equivalent matrix (derived by elementary transformations) in which the first term of jth column is $r(\lambda)$. This may be shifted by elementary transformation 1 to the leading position. It is of lower degree than $f(\lambda)$ since it is the remainder when we divided $f_1(\lambda)$ by $f(\lambda)$.

Case 2. By a similar process the same result may be achieved when the element not divisible by $f(\lambda)$ is in the first column, working only with columns instead of rows.

Case 3: When all elements in first row, first column, are divisible by $f(\lambda)$, but an element in ith row jth column is not divisible by this $f(\lambda)$. Let $k(\lambda)f(\lambda)$ represent first element in jth column. Now add to elements of jth column $-k(\lambda)$ times elements of first column. The result will be an equivalent λ-matrix (obtained by elementary transformation 3) in which the first element of the jth column is 0, and the element in ith row, jth column is still not divisible by $f(\lambda)$, and first term in ith row is still divisible. Now add to elements of first column, elements of this new jth column. We get a matrix whose leading element is still $f(\lambda)$ but the element in first of ith row is not divisible by $f(\lambda)$, so we have now a matrix of the form of case 2 which may be transformed as was that. Thus it is always possible to secure a leading element not zero of lower degree than $f(\lambda)$.

Preliminary 2: If we have given a λ-matrix whose elements are not all zero, an equivalent matrix can be formed such that the leading element $f(\lambda)$ is not zero; all the other elements of both the first row and the first column are zero; and every element outside the first row and first column shall be divisible by $f(\lambda)$.

Proof: Since the elements are not all zero (by hypothesis) we may by a sufficient number of changes of rows, and of columns, place in the leading position an element $\neq 0$. If this element $f(\lambda)$ in leading position is not a factor of all the other elements, by preliminary 1, just proved, we form an equivalent matrix whose leading element is not zero and of lower degree than $f(\lambda)$. We may continue this process until after a finite number of changes we reach a an equivalent λ-matrix whose leading element is a factor of all the others. After this we may, by using elementary transformation(3)a sufficient number of times, secure an equivalent λ-matrix in which the leading element is not zero, but all

the other elements of the first row and of the first column are zero, while all the rest are divisible by the leading element.

We are now to prove

Theorem 2: Every matrix of the nth order and of rank r can be reduced by elementary transformations to the normal form

$$\begin{Vmatrix} E_1(\lambda) & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & E_2(\lambda) & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & E_n(\lambda) & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & 0 \end{Vmatrix}$$

where the coefficient of highest power of λ in each of the polynomials $E_i(\lambda)$ is unity, and $E_i(\lambda)$ is a factor of $E_{i+1}(\lambda)$ for $i = 1, 2, \cdots, r-1$

Proof: By preliminary 2, just preceding, it was shown that if $\underline{a}$ is a λ-matrix of rank $r > 0$,

$$\underline{a} = \begin{Vmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{Vmatrix}$$

$\underline{a}$ could be transformed by a finite number of elementary transformations to an equivalent matrix $\underline{b}$ of rank r such that

$$\underline{b} = \begin{Vmatrix} f_1(\lambda) & 0 & \cdots & 0 \\ 0 & b_{11} & \cdots & b_{1,N-1} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ 0 & b_{N-1,1} & \cdots & b_{N-1,N-1} \end{Vmatrix}$$

where $f_1(\lambda) \not\equiv 0$, and where $f_1(\lambda)$ is a factor of all the b's.

The matrix $\begin{Vmatrix} b_{11} & \cdots & b_{1,N-1} \\ \cdots & \cdots & \cdots \\ b_{N-1,1} & \cdots & b_{N-1,N-1} \end{Vmatrix}$ is of order (n-1). Hence its

rank must be (r-1), since rank of last written matrix was r.

So if r>1, then by a similar process this may be transformed to an equivalent matrix

$$\underline{c} = \begin{Vmatrix} f_2(\lambda) & 0 & \cdots & 0 \\ 0 & c_{11} & \cdots & c_{1,N-2} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & c_{N-2,1} & \cdots & c_{N-2,N-2} \end{Vmatrix}$$

where $f_2(\lambda) \not\equiv 0$, and where $f_2(\lambda)$ is a factor of all the c's!

Theorem 1 of this section proved that if $f_2(\lambda)$ is the greatest common divisor of all the elements of C, it must be a ~~factor~~ greatest common divisor of the b's and so is divisible by $f_1(\lambda)$. By a series of changes then we have transformed $\underline{a}$ so that it can now be written in the form.

$$\underline{a} = \begin{Vmatrix} f_1(\lambda) & 0 & 0 & \cdots & 0 \\ 0 & f_2(\lambda) & 0 & \cdots & 0 \\ 0 & 0 & c_{11} & \cdots & c_{1,N-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & c_{N-2,1} & \cdots & c_{N-2,N-2} \end{Vmatrix}$$

where $f_1(\lambda) \not\equiv 0$, $f_2(\lambda) \not\equiv 0$, f_1 is a factor of f_2, and f_2 is a factor of the c's.

If r>2 we may treat (N-2)rowed matrix of c's in same manner, as it is of rank (r-2). Thus if we continue, we may finally transform the original matrix thus:

$$\underline{a} = \begin{Vmatrix} f_1(\lambda) & 0 & \cdots & 0 & \cdots & 0 \\ 0 & f_2(\lambda) & \cdots & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & f_r(\lambda) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{Vmatrix}$$

if f's $\not\equiv 0$ and each is a factor of each of the following f's.

By means of transformation 2, the coefficient of the highest power of λ can be made unity, and our matrix may be written in the form given in the statement of the theorem.

Theorem 3: The greatest common divisor of the i-rowed determinants of a λ-matrix of rank r ($i \leqq r$) is

$D_i(\lambda) = E_1(\lambda)\ E_2(\lambda) - - - - E_i(\lambda)$ where the E's are the elements of the normal form to which the given matrix is equivalent.

Proof: Since the original matrix a is equivalent to the normal form of the matrix, (just proven), the greatest common divisor of the i-rowed determinants of one must be the greatest common divisor of the i-rowed determinants of the other by theorem 1. But all the i-rowed determinants of the normal form are zero, except those made up of E's - for all the other elements are zero. Having thus prepared the way we may now prove the important

Theorem 4: A necessary and sufficient condition for the equivalence of two λ-matrices of the nth order is that they have the same rank r and that for every value of i from 1 to r inclusive, the i-rowed determinants of one matrix have the same greatest common divisor as the i-rowed determinants of the other.

Necessary case: This is proved by **1**, VII which states that if two λ-matrices of rank r are equivalent, the greatest common divisor of the i-rowed determinants of one is also the greatest common divisor of the i-rowed determinants of the second.

Sufficient condition:

Each matrix may be reduced to the normal form:

$$\underline{a} = \begin{Vmatrix} E_1(\lambda) & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & E_2(\lambda) & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & E_n(\lambda) & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{Vmatrix} \qquad \underline{b} = \begin{Vmatrix} E'_1(\lambda) & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & E'_2(\lambda) & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & E'_n(\lambda) & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{Vmatrix}$$

The greatest common divisors of i-rowed determinants, ($i \leq r$) of $\underline{a}$ and $\underline{b}$ are, respectively, by theorem 3 - $E_1(\lambda)\ E_2(\lambda) - - - E_i(\lambda)$

and $E'_1(\lambda)\ E'_2(\lambda) - - - E'_i(\lambda)$.

But these are equal by theorem 1 and by hypothesis. Hence $E_i(\lambda) = E'_i(\lambda)$ and the two normal forms are identical. ($i = 1, 2 - - n$) Therefore the λ-matrices from which these were derived are equivalent, since they have been shown to be equivalent to a third matrix.

Summary

In this discussion of matrices we have found that the idea of the matrix notation is a comparatively recent discovery in the field of mathematics, having been brought into being by Cayley in 1843. Frobenius, Krönecker and Weierstrass extended the theory, finding it a valuable tool in their work on invariants, linear equations, linear transformations, bi-linear and quadratic forms, quaternions and vectors. It has been shown that although matrices are not determinants, they depend upon the latter so consistently for the establishment of the necessary theorems, that a considerable number of the definitions and theorems of determinants were needed as postulates on which to build our theory. Accordingly such a list was given in section II.

In the main body of the paper the nature and properties of a matrix, together with the related theorems were discussed. As a whole, matrices were shown to fall under one of these three classes:
1) Ordinary matrices - a system of m n quantities arranged in a rectangular array of M rows and N columns; - 2) Symmetric matrices in which the pairs of equal terms are symmetrically arranged with respect to the leading diagonal; and 3) λ matrices whose elements are polynomials in a single variable λ.

The theorems developed about these main headings may be grouped as follows:

A) - Definitions and theorems which determine the rank of a matrix.

B) - Definitions and theorems which have to do with the equivalence of two matrices.

C) - Definitions and theorems concerning the conformation of the laws of matrices with the ordinary rules of algebra.

Bibliography

I An Introduction to Higher Algebra,

by Maxime Bocher. 321 pages.

1929 - The Macmillan Co., New York.

This paper is based almost wholly on the work found in pages 20-22; 54-66; 74-93; 282-283; 296-302 of this text. It is, apparently, the only development of the topic of matrices which does not demand an intensive knowledge of elementary divisors, quadratic forms, invariants and linear transformations. While it does give a very logical development of these topics, the work on matrices is developed independently and then applied as an aid in the discussion of these subjects.

II Matrices and Determinoids.

by C. E. Cullis - 3 Volumes.

1913 - 1918 - 1925 -- The University Press, London.

This is by far the most elaborately developed and detailed work which I discovered on the topic of matrices. The volumes are large, finely printed and give a very comprehensive development of the topic. The work of Dr. Bocher is quoted as the leading text by Dr. Cullis in his bibliography, and he follows in

many instances Dr. Bocher's volume just quoted. It would be impossible in a paper of this length to approach the topic as taken up by Dr. Cullis. His third volume, however, gives a very complete summary of the application of matrices in modern mathematics.

III Modern Algebraic Theories,
by Leonard Eugene Dickson.
1926 - The Benjamin H. Sanborn Co.
Chapters III to V, pages 89-134, inclusive, give an admirably brief yet comprehensive view of matrices. This development is, however, so interwoven with the theories of invariants and quadratic forms, (introduced by the author in the previous chapters), that it is impossible to develop the theory of matrices independently of these. Their introduction would demand another paper of explanation of even greater length than that of the present topic.

IV A Treatise on the Theory of Determinants.
by Thomas Muir. - 3 Volumes.
1882 - Macmillan Co., London.

Useful as a basis of preliminary work on determinants.

W. A Treatise on Universal Algebra: with applications.
by Alfred North Whitehead.
1898 - Cambridge University Press.
C. J. Clay & Sons, London.
Book IV, Chapter VI, pages 248-269.

A brief and scholarly development, but with a special terminology, and too advanced for ordinary application. As arranged by Dr. Whitehead, the chapter is a part of the Calculus of Extensions.

BOSTON UNIVERSITY
1 1719 02573 0856